

# Patient-Oriented Cancer Information on the Internet: A Comparison of Wikipedia and a Professionally Maintained Database

By Malolan S. Rajagopalan, MD, Vineet K. Khanna, MD, Yaacov Leiter, MS, MD, Meghan Stott, MD, Timothy N. Showalter, MD, Adam P. Dicker, MD, PhD, and Yaacov R. Lawrence, MA, MBBS, MRCP

Department of Radiation Oncology, University of Pittsburgh Cancer Institute; University of Pittsburgh School of Medicine, Pittsburgh; Drexel University College of Medicine; Department of Radiation Oncology, Jefferson Medical College of Thomas Jefferson University, Kimmel Cancer Center, Philadelphia, PA; Bruce and Ruth Rappaport Faculty of Medicine, Technion-Israel Institute of Technology, Haifa, Israel

## Abstract

**Purpose:** A wiki is a collaborative Web site, such as Wikipedia, that can be freely edited. Because of a wiki's lack of formal editorial control, we hypothesized that the content would be less complete and accurate than that of a professional peer-reviewed Web site. In this study, the coverage, accuracy, and readability of cancer information on Wikipedia were compared with those of the patient-orientated National Cancer Institute's Physician Data Query (PDQ) comprehensive cancer database.

**Methods:** For each of 10 cancer types, medically trained personnel scored PDQ and Wikipedia articles for accuracy and presentation of controversies by using an appraisal form. Reliability was assessed by using interobserver variability and test-retest reproducibility. Readability was calculated from word and sentence length.

**Results:** Evaluators were able to rapidly assess articles (18 minutes/article), with a test-retest reliability of 0.71 and interobserver variability of 0.53. For both Web sites, inaccuracies were rare, less than 2% of information examined. PDQ was significantly more readable than Wikipedia: Flesch-Kincaid grade level 9.6 versus 14.1. There was no difference in depth of coverage between PDQ and Wikipedia (29.9, 34.2, respectively; maximum possible score 72). Controversial aspects of cancer care were relatively poorly discussed in both resources (2.9 and 6.1 for PDQ and Wikipedia, respectively, NS; maximum possible score 18). A planned subanalysis comparing common and uncommon cancers demonstrated no difference.

**Conclusion:** Although the wiki resource had similar accuracy and depth as the professionally edited database, it was significantly less readable. Further research is required to assess how this influences patients' understanding and retention.

## Introduction

Over the past decade, the Internet has become an important source of cancer information. At one cancer center in the United States, 80% of patients had access to the Internet and 63% searched for cancer information online.<sup>1</sup> Similar findings have been described in other medical specialties.<sup>2-4</sup>

A decade ago, most Web sites were static; users passively viewed content but were unable to create or share their own ideas and knowledge. More recently, Web sites and online communities have adopted a user-centric design that encourages collaboration and personal interaction, an approach described by the term "Web 2.0."<sup>5</sup> Examples of Web 2.0 Web sites include online blogs, social networking sites (eg, Facebook) and wikis. A wiki is a Web site that allows its users to edit its pages online, creating and modifying the information that is publicly available. Among wikis, Wikipedia ([www.wikipedia.org](http://www.wikipedia.org)), an online collaborative encyclopedia, has achieved enormous popularity and has become one of the world's most trafficked Web sites, with more than 3 million English-language articles.<sup>6-7</sup>

The freely editable nature of Wikipedia enables contributors, lay or expert, across the world to share their knowledge easily. In addition, the content on the Web site can change constantly and quickly to reflect the latest news and data, a feature difficult to implement in a formally edited encyclopedia.<sup>8</sup> Unfortunately, such a system can also propagate misinfor-

mation if users contribute erroneous data either accidentally or maliciously. Although some such examples have become notorious, they appear to be uncommon.<sup>8</sup> Because there is no formal editing process, such errors will be corrected only if they are recognized and deleted by other users.

The quality and accuracy of medical content available online is a matter of concern. In particular, are user-edited wikis providing accurate, balanced, comprehensive information to patients and the wider community? We were especially concerned that wikis may misconstrue nonscientific proven beliefs as fact or contain outright mistakes. Conversely, we hypothesized that the wiki platform may be more conducive to a discussion of controversial aspects of cancer care. To address these issues, we compared the cancer information available on Wikipedia with that of a more traditional online source, the patient-oriented Physician Data Query (PDQ; [www.cancer.gov/cancertopics/pdq](http://www.cancer.gov/cancertopics/pdq)) maintained by the National Cancer Institute (NCI). The PDQ is a series of professionally edited and maintained articles that serve to educate patients with cancer about their disease. In contrast to Wikipedia, the PDQ is not available for online editing or modification. It is produced by editorial boards composed of oncologists, psychologists, geneticists, epidemiologists, and complementary medical practitioners according to a rigid process that includes data gathering, writing, and editing.<sup>9</sup> This study sought to compare these two resources across a variety

of domains, including accuracy of content, depth of coverage of topics, presentation of controversial issues in care, and readability.

## Methods

### Overall Study Design

Appraisal forms assessing the depth and accuracy of information and the presentation of controversies were produced for each cancer type. In order to ensure that all evaluators were using the same text, PDFs were created from the PDQ articles and Wikipedia articles as they appeared on August 26, 2009. In cases where there were hyperlinks to highly relevant subarticles, these were included as well. For the PDQ resource, these links included prevention and screening articles. For Wikipedia, major linked subheadings (represented on the parent article as “see also”) were provided to the evaluators. These included articles on staging, screening, and epidemiology, among others, and varied on the basis of cancer type. The evaluators were instructed to score only the material presented to them in PDF format, and not to consult with the Internet.

### Cancer Sites

Five of the most common cancers and five of the less common cancers were chosen on the basis of statistics published by American Cancer Society.<sup>10</sup> The common cancer types were lung, breast, prostate, colon, and melanoma, which have an annual incidence in the United States ranging from 68,720 to 219,440. The less common cancer types were anal, vulvar, small intestine, testicular, and osteosarcoma; these cancer sites have an annual incidence ranging from 2,570 to 8,400.<sup>10</sup>

### Assessment of Accuracy and Depth of Coverage

For each cancer type, the study designers (M.S.R., Y.R.L.) assembled an appraisal form that contained eight information statements obtained from *Abeloff's Clinical Oncology* textbook<sup>11</sup> and validated in Devita's *Principles and Practice of Oncology*.<sup>12</sup> Information statements, which were likely to be of interest and relevance to patients, were chosen, encompassing the domains of epidemiology, etiology, symptoms, diagnosis, and treatment. Evaluators were instructed to compare each information statement with the content in the Wikipedia and PDQ PDFs and score the accuracy of the information presented. The scoring system was as follows: 0 points if the topic was not discussed, 1 point if there was a discussion of the topic but with major omissions, 2 points if there was discussion of the topic with only minor omissions, and 3 points if there was a complete discussion of the topic; -1 point was scored if the content in the resource was discordant with the information statement. In addition, we examined the incidence of errors. An error was recorded only when two or more evaluators agreed that a Web site presented a particular piece of information that was discordant with the appraisal form.

### Assessment of Controversial Aspects of Care

Two controversial topics for each cancer type were assessed. These topics were chosen on the basis of a literature search using the terms “disputed,” “controversial,” or “disagreement.” The

scoring system was as follows: 0 points if the topic was not mentioned, 1 point if the topic was raised but the controversial aspect was avoided, 2 points if the controversy was briefly discussed, and 3 points for complete discussion of the controversy.

One month after completion of the study, in order to assess the test-retest reliability, we asked all the evaluators to re-evaluate both the PDQ and Wikipedia articles for one cancer site, using the same PDFs previously used. At the conclusion of the assignment, evaluators were questioned regarding the time needed to assess each web resource and to provide additional feedback and criticism.

### Assessment of Readability

Readability was calculated by using the validated Flesch-Kincaid grade level, which factors both word choice (number of letters per word) and sentence structure (number of words per sentence).<sup>13</sup> For each article, three passages of text ranging between 90 and 120 words in length were randomly selected from the beginning, middle, and end. This text was edited to remove any citations, titles, headings, and references so as not to artificially alter the calculated grade level. The Flesch-Kincaid grade level was calculated by using Word 2007 (Microsoft, Redmond, WA), and the average readability score for each article was determined.

### Assessment of References

For each article, the number of references was noted. In addition, each reference was categorized into one of the following groups: academic journals, books, professional organizations, news media (print or web), commercial (including pharmaceutical companies), and other. A subanalysis comparing these results between common and uncommon cancers was also performed.

### Frequency of Revisions to Wikipedia Articles

The Wikipedia Web site documents each time a particular article is changed or edited by its users. For the 10 cancer types analyzed, we counted the number of edits over the course of the year preceding this study (August 27, 2008 through August 26, 2009). The number of edits was averaged, and comparisons between common and uncommon cancers were performed.

### New Information

The ability of the Web sites to integrate new information was compared by assessing whether they contained the results of recently published clinical trials. We noted whether the relevant Web sites referred to the results of the 10 most recent clinical trial manuscripts of solid tumors in adults published in the *New England Journal of Medicine* (April 2011 to February 2011). This assessment was performed in mid-February 2011.

### Statistical Analyses

Comparisons between the Web sites were made by means of a paired two-sided *t* test, with significance set at  $P < .05$ . Reliability was assessed by using interobserver variability (correlation coefficient) and test-retest reproducibility (Ebel's algorithm).<sup>14</sup>

## Results

### Descriptive Comparison of the Two Resources

Wikipedia is an online collaborative encyclopedia that relies on millions of users to create and edit its articles. With a staff of only 35, the Wikimedia Foundation that oversees Wikipedia has no role in content oversight. Articles are updated at any time by any of its users. In contrast, the NCI PDQ has an editorial board composed of experts in the fields of medical, radiation, and surgical oncology. The editorial board reviews published research studies on a monthly basis and meets eight times each year to consider new information to integrate into the articles.

Both resources have a fairly consistent organizational scheme for articles. Wikipedia articles begin with a short introduction followed by a table of contents. Most articles then include sections on signs and symptoms, causes, pathogenesis, diagnosis, screening, treatment, and prognosis. Articles concerning uncommon cancers lacked some sections; conversely, articles about common cancers often included additional sections such as history, prevention, and details of staging. All Wikipedia articles contain external links and references (clearly annotated in the text). PDQ articles were found to have the following sections: general information (including overview, risk factors, signs/symptoms, diagnosis), staging, and treatment options. Separate articles on prevention and screening are also included. Neither external links nor references are provided.

Recent changes to both resources are documented. The PDQ includes a last-modified date along with a short sentence describing the nature of the changes. In contrast, Wikipedia articles have a link detailing every change made to the document since its creation.

### Overview of Reviewers

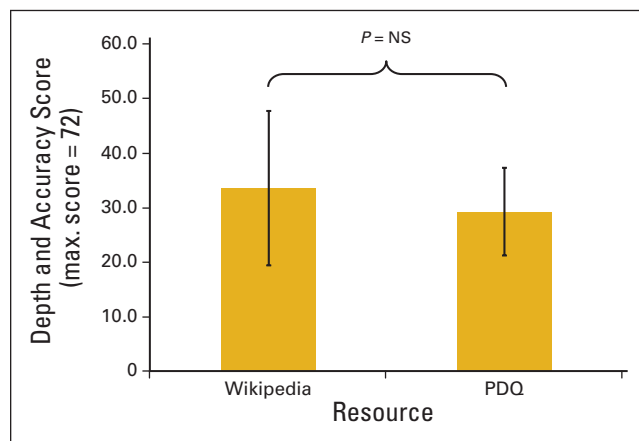
Evaluators required an average of 18 minutes to assess each article. Test-retest reliability (1 month later) was calculated to be 0.71. Using Ebel's algorithm, the interobserver variability was found to be 0.53 (a value of 1.00 would indicate perfect reliability).

### Depth and Accuracy of Content

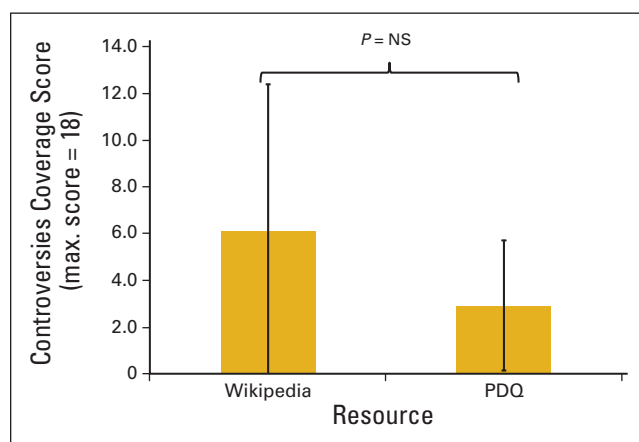
The maximum possible score for content for each resource was 72. There was no difference in the combined depth and accuracy of content between the Web sites ( $29.9 \pm 8.3$  standard deviation [SD],  $34.2 \pm 14.0$  SD for PDQ and Wikipedia, respectively) (Figure 1). Errors were found to be rare in both resources. Of the 80 information statements presented, there were zero errors in PDQ articles and one in Wikipedia (0% *v* 1%; NS).

### Coverage of Controversial Aspects of Care Was Not Different Between the Resources

The maximum possible score for complete coverage of controversial aspects averaged across the 10 disease sites was 18. Controversial aspects of cancer care were poorly discussed in both resources ( $2.9 \pm 2.8$  SD and  $6.1 \pm 6.3$  SD for PDQ and Wikipedia, respectively; NS) (Figure 2).



**Figure 1.** Depth and accuracy of content. Information statements obtained from professional oncology texts were compared against content from both electronic resources and scored. max., maximum; PDQ, Physician Data Query.



**Figure 2.** Coverage of controversial topics. Topics of controversy or debate were generated for each cancer site as described in the Methods section and were compared between the electronic resources. max., maximum; PDQ, Physician Data Query.

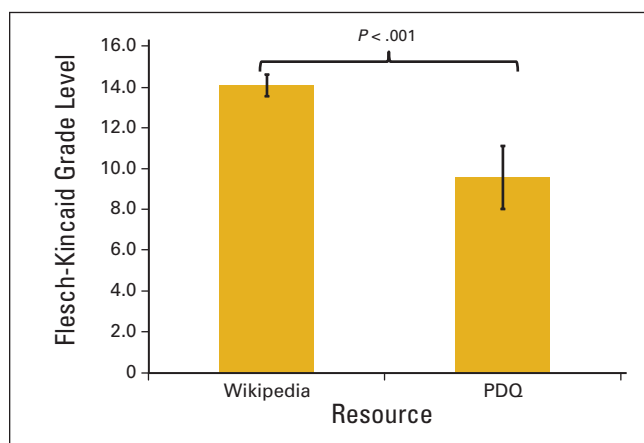
### Readability

PDQ articles were found to be significantly more readable than those on Wikipedia, with a grade level of  $9.6 \pm 1.5$  SD versus  $14.1 \pm 0.5$  SD ( $P < .001$ ) (Figure 3). This difference in grade level between PDQ and Wikipedia was preserved when analyzing common cancers ( $8.5 \pm 1.2$  SD *v*  $13.9 \pm 0.4$  SD,  $P < .001$ ) and uncommon cancers ( $10.7 \pm 0.9$  SD *v*  $14.3 \pm 0.6$  SD,  $P < .001$ ) individually.

In addition to this disparity in crude readability, we were concerned that there might be an additional difference regarding the use and explanation of technical words. Both Web sites make extensive use of hypertext links, with multiple links per paragraph. An important distinction is that whereas the PDQ hypertext linked to a dictionary written in plain English, the Wikipedia hypertext most often linked to highly technical articles.

### References

Common cancers had significantly more citations than uncommon cancers ( $98.6 \pm 33.9$  SD *v*  $10.0 \pm 6.5$  SD;  $P = .0036$ ).



**Figure 3.** Readability. Readability was assessed by using the Flesch-Kincaid grade level, factoring word length and sentence structure. PDQ, Physician Data Query.

Common cancer articles were found to have a significantly higher percentage of Medline citations (62.2% *v* 24.2%;  $P = .037$ ) and a lower percentage of citations from nonprofit organization or foundation Web sites (20.7% *v* 48.7%;  $P = .020$ ). No differences were detected in percentage of citations from books, news media, commercial, and other (Appendix Figure A1, online only). The patient-oriented PDQ does not include references.

### Frequency of Revisions to Wikipedia Articles

We compared the frequency of revisions to Wikipedia articles about common cancers with those about uncommon cancers (Appendix Figure A2, online only). Articles about uncommon cancers were edited significantly less frequently than those about common cancers ( $115.4 \pm 113.2$  SD *v*  $513.8 \pm 216.9$  SD;  $P = .011$ ).

### Evaluator Feedback

Feedback obtained at the conclusion of the study revealed that appraisal of an article required an average of 18 minutes. In addition, the evaluators were asked to rate the fairness of the appraisal form on a scale of 1 to 5 (with 5 representing the most fair); average score was 4.7.

### New Information

A study was performed to assess the integration of current data within each resource. Out of a maximal possible score of 10, Wikipedia scored 4 and PDQ 0 ( $P < .04$ ).

### Additional Studies

In order to put our results in context, we investigated which Web site was favored by popular search engines. A variety of cancers were searched for by using both Google (www.google.com) and Bing (www.bing.com). Both Wikipedia and PDQ links typically appeared within the top 10 search results. In more than 80% of cases, Wikipedia appeared above PDQ in the results list.

## Discussion

Patients and their families are increasingly turning to the Web as a source of medical information.<sup>15</sup> In the last several years, there has been a tremendous growth in collaborative, Web 2.0 sites such as the online public encyclopedia Wikipedia, which enables any visitor to contribute and edit any article (although certain politically sensitive articles are not amenable to open editing, this does not apply to cancer-related information). In this study, we sought to determine how Wikipedia articles about five common and five uncommon cancers compared with articles about the same cancers from a peer-reviewed, expert-generated Web site, NCI's PDQ. The domains used for this comparison included depth of content, accuracy, discussion of controversial topics, and readability. In addition, the type and quality of references cited in Wikipedia articles, as well as the frequency with which these articles were edited, was assessed.

We found that although Wikipedia had similar accuracy and depth to the PDQ, the written style was more complex and thus might be less understandable to patients. We found no difference in the discussion of controversial topics between the two resources. Although the Wikipedia articles appeared to be more up-to-date, we acknowledge that this may possibly reflect a policy of PDQ not to discuss published studies until the pharmaceutical agents have been approved by the US Food and Drug Administration. Finally, regarding references in Wikipedia, more common cancer types had significantly more citations, as well as a higher proportion of citations from Medline-indexed articles.

To our knowledge, this is the most comprehensive and rigorous study comparing oncology articles between an expert-generated Web site and a wiki. Many parameters that are directly relevant to a patient's perspective, including accuracy, depth, and readability, were assessed. Weaknesses of the study include that its scope was limited to five common and five uncommon cancers, the fact that only a limited number of Web sites were examined, and the use of medically trained evaluators. In future studies, we intend to use a larger number of evaluators who are more representative of the general population.

Studies such as this one that assess the quality of online content will be ever more important as the Internet continues to increase in influence.<sup>16</sup> Wikipedia is an especially prominent Web site; on general online searches for various medical terms and diseases, Wikipedia articles ranked among the first 10 results for 71% to 85% of the search engines and key words tested, surpassing professionally maintained Web sites (eg, National Institutes of Health MedlinePlus and National Health Service Direct Online).<sup>17</sup> There have been very few studies assessing the quality of Wikipedia's medical articles. One study that focused on osteosarcoma found that Wikipedia was inferior to the NCI Web site.<sup>18</sup> Another study that assessed the description of surgical procedures by Wikipedia found that although all the entries presented accurate content, 37.1% of articles had at least one critical omission. Interestingly, the study found a positive correlation between the frequency with which an article was edited and its accuracy.<sup>19</sup> The latter finding partially concurs with our finding that articles about more com-



mon cancer types, which were significantly more frequently edited, had better quality references than those about uncommon cancers.

Our data indicate that although the informational content of articles on Wikipedia and PDQ is comparable, the former are less easy to read. The hypertext links provided by PDQ to a lay language dictionary further promote understanding, although this was not formally assessed in the readability metric. The implications of this disparity in reading grade are not known. Several studies have concluded that those patients who look for information online have above-average educations.<sup>1,17,20</sup> However, many patients with cancer have impaired cognitive function.<sup>21</sup> A complete understanding of the impact of disease, treatment, and educational attainment on the understanding and retention of Web-based information, although important, was beyond the scope of this study.

In conclusion, we found that Wikipedia and PDQ entries have comparable depth and accuracy, but the former were significantly less readable. On the basis of the sample articles tested, both appear to be reliable sources of information, but the editorial processes used by PDQ created a more readable result. Further research is required to ascertain what patient- and Web page-related factors determine optimal understanding and absorption of information. Such research will help in the design of the next generation of Web-based information systems.

*Accepted for publication on March 24, 2011.*

#### Acknowledgment

The Kimmel Cancer Center (A.P.D., Y.R.L., T.S.) is supported by National Cancer Institute Grant No. 2 P30 CA056036-09. Y.R.L. is sup-

ported by a Young Investigator Award from the American Society for Clinical Oncology.

Presented in part at the 46th Annual Meeting of the American Society of Clinical Oncology, June 4-8, 2010.

#### Authors' Disclosures of Potential Conflicts of Interest

The authors indicated no potential conflicts of interest.

#### Author Contributions

**Conception and design:** Malolan S. Rajagopalan, Adam P. Dicker, Yaacov R. Lawrence

**Administrative support:** Malolan S. Rajagopalan, Adam P. Dicker, Yaacov R. Lawrence

**Provision of study materials or patients:** Malolan S. Rajagopalan, Timothy N. Showalter

**Collection and assembly of data:** Malolan S. Rajagopalan, Vineet K. Khanna, Yaacov Leiter, Meghan Stott, Timothy N. Showalter, Yaacov R. Lawrence

**Data analysis and interpretation:** Malolan S. Rajagopalan, Vineet K. Khanna, Yaacov Leiter, Adam P. Dicker, Yaacov R. Lawrence

**Manuscript writing:** Malolan S. Rajagopalan, Vineet K. Khanna, Yaacov Leiter, Adam P. Dicker, Yaacov R. Lawrence

**Final approval of manuscript:** Malolan S. Rajagopalan, Vineet K. Khanna, Yaacov Leiter, Meghan Stott, Timothy N. Showalter, Adam P. Dicker, Yaacov R. Lawrence

Corresponding author: Yaacov R. Lawrence, MBA, MBBS, MRCP, Department of Radiation Oncology, Sheba Medical Center, Tel HaShomer 52621, Israel; e-mail: yaacov.lawrence@sheba.health.gov.il.

DOI: 10.1200/JOP.2010.000209; posted online ahead of print at <http://jop.ascopubs.org> on August 4, 2011.

## References

- Castleton K, Fong T, Wang-Gillam A, et al: A survey of Internet utilization among patients with cancer. *Support Care Cancer* 18:18, 2010
- Schwartz KL, Roe T, Northrup J, et al: Family medicine patients' use of the Internet for health information: A MetroNet study. *J Am Board Fam Med* 19:39-45, 2006
- Diaz JA, Griffith RA, Ng JJ, et al: Patients' use of the Internet for medical information. *J Gen Intern Med* 17:180-185, 2002
- McMullan M: Patients using the Internet to obtain health information: How this affects the patient-health professional relationship. *Patient Educ Couns* 63:24-28, 2006
- Giustini D: How Web 2.0 is changing medicine. *BMJ* 333:1283-1284, 2006
- comScore: comScore Media Metrix Ranks Top 50 U.S. Web Properties for May 2010, June 15, 2010. <http://ir.comscore.com/releasedetail.cfm?ReleaseID=479764>
- Wikipedia: About - Wikipedia, July 23, 2010. <http://en.wikipedia.org/wiki/Wikipedia:About>
- Giles J: Internet encyclopaedias go head to head. *Nature* 438:900-901, 2005
- National Cancer Institute: PDQ - NCI's Comprehensive Cancer Database, January 4, 2010. <http://www.cancer.gov/cancertopics/pdq/cancerdatabase>
- Jemal A, Siegel R, Ward E, et al: Cancer statistics, 2009. *CA Cancer J Clin* 59:225-249, 2009
- Abeloff MD, Armitage JO, Niederhuber JE, et al: *Abeloff's Clinical Oncology*. Philadelphia, Churchill Livingstone, 2008
- DeVita VT, Lawrence TS, Rosenberg SA, et al: *DeVita, Hellman, and Rosenberg's Cancer: Principles & Practice of Oncology*. Philadelphia, Lippincott Williams & Wilkins, 2008
- Kincaid JP, Fishburne RP, Rogers RL, et al: Derivation of New Readability Formulas (Automated Readability Index, Fog Count, and Flesch Reading Ease Formula) for Navy Enlisted Personnel. Research Branch Report, Memphis Naval Air Station 8, 1975
- Solomon DJ: The rating reliability calculator. *BMC Med Res Methodol* 4:11, 2004
- Huang GJ, Penson DF: Internet health resources and the cancer patient. *Cancer Invest* 26:202-207, 2008
- Cima RR, Anderson KJ, Larson DW, et al: Internet use by patients in an inflammatory bowel disease specialty clinic. *Inflamm Bowel Dis* 13:1266-1270, 2007
- Laurent MR, Vickers TJ: Seeking health information online: Does Wikipedia matter? *J Am Med Inform Assoc* 16:471-479, 2009
- Leithner A, Maurer-Ertl W, Glehr M, et al: Wikipedia and osteosarcoma: A trustworthy patients' information? *J Am Med Inform Assoc* 17:373-374, 2010
- Devgan L, Powe P, Blakey B, et al: Wiki-surgery? Internal validity of Wikipedia as a medical and surgical reference. *J Am Coll Surg* 205:S76-S77, 2007
- Baker L, Wagner TH, Singer S, et al: Use of the Internet and e-mail for health care information: Results from a national survey. *JAMA* 289:2400-2406, 2003
- Wefel JS, Witgert ME, Meyers CA: Neuropsychological sequelae of non-central nervous system cancer and cancer therapy. *Neuropsychol Rev* 18:121-131, 2008

